DOCUMENT RESUME

ED 347 120 SO 030 194

AUTHOR Torney-Purta, Judith
TITLE Measuring Performance in Social Studies in an Authentic Fashion.
PUB DATE 90
NOTE 21p.; Paper presented at the Annual Meeting of the American Educational Research Association (Boston, MA, April 1990).
PUB TYPE Reports - Research/Technical (143) -- Speeches/Conference Papers (150)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS *Cognitive Processes; Educational Research; Elementary Secondary Education; *Evaluation Methods; Political Socialization; Schemata (Cognition); *Social Studies; Student Development; *Student Evaluation; Testing; Thinking Skills
IDENTIFIERS *Authenticity; Higher Order Learning

ABSTRACT

Researchers of performance assessment in social studies have contrasted authentic assessment, which is characterized by contextualized, complex intellectual challenges, with multiple choice tests of achievement, which are characterized as fragmented and static. In calling for the need for authentic assessment, this paper outlines a plan for the development of performance indicators in social studies, in particular, indicators of complex cognitive representations of social institutions and structures. Four characteristics of complexity in the international political and economic views of young people are identified based on research consisting of problem-solving interviews and a written questionnaire pre- and post-simulation. The ideas put forth are not meant to provide an immediate alternative tc multiple choice testing in the social studies, but rather represent a starting point by defining what is meant by complexity of cognitive structures with respect to economic and political systems. A list of references is included, as are several figures depicting the schemata of student research subjects. (DB)

Measuring Performance in Social Studies In an Authentic Fashion

Judith Torney-Purta

Professor of Human Development

University of Maryland at College Park

College Park, MD 20742

Measuring Performance in Social Studies In an Authentic Fashion

Judith Torney-Purta
Professor of Human Development
University of Maryland at College Park
college Park, MD 20742

Paper presented at symposium on Performance Assessment, American Educational Research Association, Boston, April, 1990. Some portions of this paper will appear in a chapter in the volume reporting on a conference on Education Quality Indicators held at UCLA in October, 1989. Copyright to Figures 1 and 2 and their descriptions is held by Teachers College Press, publixher of Torney-Purta (1990). The support of the center for Educational Research and Development at the University of Maryland is gratefullyu acknowledged.

Performance assessment in social studies and its related fields (history, geography, civics and government) is in a position similar to that of performance assessment in science five years ago. This paper will outline a plan for the development of indicators of performance in social studies, in particular indicators of complex cognitive representations of social institutions and structures in relation to the need for authentic assessment.

Wiggins (1989) has recently contrasted authentic tests, which he characterizes as "contextualized, complex intellectual challenges" with "fragmented and static" multiple choice measures. Authentic measures are not restricted to recall and do not reflect lucky or unlucky one-shot responses. He has provided four criteria of authenticity in testing, three of which will be addressed in relation to social studies performance assessment:

> Intellectual design features of an authentic test emphasize realistic (but fair) complexity, stress depth more than breadth, include some ill-structured tasks or problems which the student must structure, and require students to contextualize the content knowledge which they have acquired.
>
> Standards of grading and scoring features of an authentic test include complex multifaceted criteria which can be specified and reliably agreed upon by scorers. What constitutes an authentically high level of performance can be explained to students and to teachers (and can serve as a guide to instruction that makes testing a central experience in learning).
>
> Fairness and equity features of an authentic test allow students to show off what they can do and accommodate individual student's interests and strengths. They strike a balance between "honoring achievement, progress, native language skill, and prior fortunate training." Wiggins, 1989, pp. 711-12)

In delineating each of these criteria Wiggins refers with disparagement to multiple choice tests of achievement restricted to the recall of facts. Items too often find their way into a test because they show good distributions (with 40 to 60 percent of students answering correctly) and contribute to a total score which can be graphed as a normal curve. In fact, authentic indicators of performance are superior to multiple choice tests on nearly every index except cost and ease of administration and scoring.

A similar stress on the importance of developing new methods

is discussed in the field of psychological tests, especially intelligence tests, by Miller-Jones (1989). This author call for the development of instruments which allow inferences about cognitive processes from test performance. Intelligence tests are criticized because in addition to the normal curve criterion, answers are often designated correct primarily because they are answers more frequently given by older children in the normative sample. Miller-Jones is especially concerned about the equity issue and the failure of many current testing practices to capture the "socio-cognitive ecology" of minority group children who have access to cognitive representatives which are adaptive but which differ from those of the majority group.

Let us move from generalized statements attacking multiple choice and standardized tests to a consideration of assessments in the social studies. The correct answers to multiple choice tests in fields such as history, civics, and geography usually rely on recognition of discrete facts and definitions. Little attempt is made to construct items for these tests around the cognitive structures or key ideas to which young people might relate meaningful knowledge about social or political institutions. Compounding this lack of attention to key ideas or to underlying modes of connection, is the guessing factor. The large majority of multiple choice tests in the social studies provide no way to distinguish between the student who has answered an item correctly by guessing, or by accessing a piece of information learned by rote which is unconnected to any cognitive structure, or by accessing a piece of information using a cognitive structure which has organized the necessary knowledge to answer the question (e.g. a time line, mental map, or image of the economic or political system). In multiple choice mathematics tests it is possible to construct incorrect answers which represent the application of a commonly used but wrong process of thought or calculation; it is extremely difficult to construct such informative distractors in questions about social studies or history. Norris (1989) has proposed having small groups of students think aloud while choosing answers to multiple choice questions as a way of improving the alternatives given. Although this procedure probably allow the tester to note when guessing is the basis for an answer, it does not appear to probe effectively for connections between information and cognitive structures. The time is ripe to build replacements for multiple choice social studies tests on recent advances in cognitive psychology.

The approach to performance assessment in social studies outlined in this paper is based on an important assumption which relates to the intellectual design features that Wiggins and Miller-Jones suggested. The assumption is that there are cognitive structures, in particular schemata and scripts, which are important in students' learning and subsequently accessing the factual material which is presented in social studies.

Prawat (1989) notes the importance of the organization of the knowledge base or its connectedness, as well as the importance of main ideas which serve as organizers for other pieces of information. Some learning situations do not result in the connection of a piece of factual material to a cognitive structure; that is often called "piecemeal" or unconnected storage of information. Many social studies lessons and textbooks can be characterized as dull and "inconsiderate" because factual material is presented in precisely this unconnected fashion. See for example Stodolsky's recent book (1988) presenting classroom observations during fifth grade social studies and math lessons and Armbruster's analysis of social studies textbocks (Armbruster, 1984).

There are in fact several types of structures which students might use in retaining and accessing learned material in the social studies. In history, a temporal organizing structure might be employed. Or key concepts such as nation might serve such a function. Or ways of visualizing and elaborating the sequences of actions or scripts describing the exercise of political leadership either at the national level or in one's neighborhood might organize knowledge. Geography has in fact provided a frequently used analogy for cognitive structures or schemata, the idea of "mental maps."

If one accepts the conclusion that existing tests should be replaced by methods which allow the assessment of cognitive structures, schemata, or scripts, the task becomes one of exploring various approaches, in particular those suggested by cognitive psychology, and of delineating as clearly as possible what is meant in saying that a student has a complex or comprehensive cognitive structure around which specific knowledge of a political, social, or economic system (contemporaneously or in the past; locally, nationally or internationally) is organized. A major goal of my current research is to conceptualize what understanding of social and political institutions means in terms of cognitive structures and then to assess these structures in an authentic way by measuring the elements of knowledge they contain, connections between these elements, and their constraints or limitations. This can be accomplished by inferring a student's underlying cognitive structure from what he or she says or writes in response to a hypothetical problem or a question.

Research with Adolescents: Think-Aloud Problem Solving

My research aiming to define these complex cognitive structures is based on both interviews and written short answer questions. The research site for exploring these assessments is a computer-assisted foreign policy simulation (Project ICONS) which has taken place for the last five summers as part of the Maryland Summer Center for International Studies. Forty 12- to

17-year-olds from Maryland are divided into teams representing diplomats from six nations,-- Brazil, France, Japan, Nigeria, Mexico and the USSR. Entry to the program is competitive, and many but not all of the participants are designated by their schools as gifted and talented students. A team representing the United States is played by university students. Following briefing about foreign policy, study of a scenario of foreign policy situations set in the near future, and the preparing of a position paper outlining the foreign policy aims of their country, the participants begin a negotiation process that takes place by teams sending messages using a computer-conferencing link. All within-team communication takes place face-to-face; all between-team communication takes place over the computer network, with printed and onscreen text available. Observations of the participants convinced me that there were no existing tests which would adequately measure the changes in representations of the world taking place in these students (Torney-Purta, 1989).

Some techniques used by cognitive psychologists to study the solving of "ill-structured" social problems pointed in an interesting direction. Voss and his colleagues (1983) developed methods for comparing the processes which novices and experts used in social science problem solving. These researchers used think-aloud problems such as the following: "Assume you are the head of the Soviet Ministry of Agriculture, and assume crop productivity has been low over the past several years. You now have the responsibility of increasing crop productivity. How would you go about doing this?" In analyzing these think-aloud protocols, the sequence of different elements of the argument (e.g. stating a subproblem, stating a solution, evaluating a solution, stating a fact) was charted. Voss noted that experts (professors specializing in Soviet affairs) spent much more time defining a problem and mentioned more constraints upon proposed solutions than novices (undergraduates).

Voss's method of analysis gives a relatively content-free picture of respondents' cognitive strategies and sequence of problem solving. Much less attention is given to the cognitive structures or underlying schemata which serve to organize knowledge. However, think-aloud solutions to hypothetical problems appeared to be an interesting type of performance to explore, so I constructed some problems connected to issues students were negotiating about in the Maryland simulation. These problems were presented in an interview on the first day of the two-week session and on the last day of the simulation exercise (ten days later). Much of the data presented in the remainder of this paper is based on the following interview questions:

> Problem 1: Imagine you are the finance minister of a developing country. The interest payment on your debt to

banks in the developed countries is due, but there is not enough money in your treasury to pay it. What actions would you take to solve this problem? What would you do; what would you ask others to do? (The Finance Minister's Debt Problem)

Problem 2: Imagine you are a diplomat in a country. You hear that the government of a neighboring country, let's call it Country C, is planning a system of laws very much like Apartheid which would apply to a group of immigrants who are of a different race from the others who live in Country C. Your country is very much against Apartheid. What actions would you take to solve this problem? What would you do; what would you ask others to do? (The Apartheid Problem).

I have also used a domestic problem, a mayor who hears that the property under a school has been polluted by a local company. Students have also answered some questions in a written form, described later.

Trained interviewers presented these problems to participants in this program on the second day of the two week session and again ten days later, as the simulation was concluding. After the interviewee stopped volunteering problem solutions, the interviewer would ask whether there were any problems with the solutions given, whether there were any reasons they would not work. Some respondents were also asked about their attitude toward Apartheid and toward the debt of developing countries and whether there were any solutions they had considered but not given because they thought the actions were unethical.

Research on logical problem-solving was a major source for the mode of analysis. Hayes and Simon (see Hayes, 1981) studied the processes by which an individual encodes the written instructions for a complex logical problem, looking at the "actors" and "legal" ways for them to operate. There were intriguing similarities between their discussion of actors, actions and constraints upon actions in solving logical problems and the way in which specialists in international relations discuss political actors (e.g., leaders or international organizations) and actions in which they can engage (e.g., negotiating or declaring war). Merging the identification of actors and constrained actions used to analyze solutions to logical problems with political scientists' designations of international actors and their actions led to a methodology for analyzing adolescents' responses to the two problems listed.

For the Finance Minister Problem, the interview protocols of a subset of respondents were analyzed to identify actors who might be approached by the finance minister (e.g. banks), actions

which they might undertake (e.g., refinancing the loan), constraints on actions (e.g., banks might not agree), and connections between actions (e.g., one country's defaulting might lead to problems in the economy of another country). A graphic model was drawn to represent the schema of the international system implied in each respondent's answers.

Let us first examine these graphic models for the think-aloud responses to the Finance Minister Problem from two of the Maryland adolescents (see attached figures). Pre-session responses appear on the top of the page; post responses appear on the bottom. Instead of comparing novices and experts, as in many cognitive psychology studies, adolescents' responses were compared before and after the intense and involving educational experience of role-playing an international political decision-maker.

The basic elements of the models, represented by triangles, are the actors mentioned who might be approached by the finance minister or involved to solve the problem. The most frequently mentioned actors were the banks who held the loans and the governments or economies of the countries where the loans were held. If interest payments or defaults were mentioned, it was assumed that it was the banks which were being discussed. The governments or economies of other countries were included in the figure only if they were explicitly mentioned or if the proposed action was much more likely to be undertaken by another country's government or economic system than by a bank (e.g., exporting or importing goods or giving foreign aid). On the figures, the ovals represent particular actions which these actors might perform (e.g. the banks might be asked to reschedule a debt; one's own economy might encourage domestic markets for goods). Below that in the diamonds are represented constraints on these actions, elicited by a probe question -- "are there any problems with the solutions you have suggested, any reasons they might not work?" No sequence of discussion is indicated in these figures.

On the average, more actors and more actions were mentioned after the simulation. In particular the students were more likely after the simulation to propose getting together with other Southern or debtor nations to form an interest cartel or in some way to put pressure as a group on the developed countries to lighten the debt load. They were also much more likely after the simulation to refer to actions within their own economies, particularly austerity measures and to mention organizations such as the World Bank or the IMF.

For most individuals there was some similarity between the pre- and post-session interviews regarding the actors in the economic system who were mentioned and the actions which might be asked or expected of them (compare the top and bottom of each figure). However, the schemata were clearly more complex after

the simulation experience for the majority of students whose responses have been graphed (six participants from 1987 and six from 1989). The technique seems to capture what might be called the restructuring of political/economic schemata or concept networks.

Let us examine two of the schema-maps in some detail. The schema of the international economic system for the individual whose response is graphed in Figure 1 was very rudimentary at the pre-session interview. The only actions mentioned had to do with investment in and export to another country's economy. After the simulation three additional actors were mentioned -- multi-national corporations, lending institutions in other countries, and other Southern nations. Only relatively basic constraints were given.

Figure 2 represents the respondent, a girl, who showed the greatest change in complexity of schema. There were only two actors mentioned in the pre-session interviews, banks and another country's government. Negotiation was mentioned as were several constraints and impact on the world economy. However, the language was vague ("reschedule your thing"). At the post-session interview, the banks and other developed countries' governments, were given quite complex activities depending on negotiation, recognition of reciprocity ("depends on what you are able to give") and a sense of sequential activities ("tie interest payments to economic growth"). One's own economy, a new actor, was given three possible actions ranging from diversification to austerity measures. Two of the three were seen with constraints and those were linked to the global economy. In addition, other developing countries were mentioned. The linkages or connections between potential actions were indicated with arrows.

Figures 3 and 4 present schema maps for two individuals responding to the Apartheid problem. Figure 3 shows a girl who was generally pessimistic about the possibility of doing much about Apartheid, but who nevertheless had both more actions and more constraints in her postsession responses. She was asked about possible unethical actions and responded with "start a war."

Figure 4 presents the schema map for a male from the Brazilian team. Before the simulation he mentioned three actors (your country, the other country, and your allies) each with one action. Two of the actions are constrained, that is reasons are given why they might not work. After the simulation, there are six actors, each with one or two actions. Only three constraints were mentioned, however.

These think-aloud problem solving interviews and the associated schema-maps have been used to develop a description of

what it means to say that a student has a complex cognitive structure for understanding political or economic issues. These definitions of complexity are the first step toward performance criteria for assessment indicators.

Complexity Characteristic 1: A student with a complex structure relating to a domestic or international political situation can visualize or access a variety of different solutions to a problem utilizing different actors (persons or institutions) with different actions in their repertoires.

In giving an answer to the Finance Minister problem students with a rudimentary understanding of the political or economic world limited their response to having the finance minister ask a bank to extend the loan period or cut the interest. Students possessing a more complex structure of understanding suggested other actors and actions -- e.g., the finance minister's country might undertake austerity measures, the economy might be stimulated toward more exports, or other developing countries might be asked to undertake cooperative action.

It is also possible to differentiate more complex from less complex schema-maps from responses to the problem of a diplomat facing the proposed institution of Apartheid in a neighboring country. But first another complexity characteristic:

Complexity Characteristic 2: A student with a complex structure relating to a domestic or international political problem can see constraints on the effectiveness of possible solutions.

As a follow-up to their initial response, students were asked whether they could see any problems with the solutions they had given, any reasons they might not work. Some students saw only a few relatively rudimentary constraints. For example a constraint on solutions to the Finance Minister's Debt Problem was that the banks might not agree to reduce the loans. Others not only saw a variety of constraints, but also connected them to each other -- e.g., if you got together with other developing countries to refuse to pay, you might bankrupt the banks and that would have an impact on the world economy. Constraints in the Apartheid problem included the difficulty of changing racial attitudes and associated policies and the possibility that economic sanctions, if enacted, would unintentionally hurt the racial minority. The ability to see the complexity of problem solutions and not merely to think that requests for debt aid will always be answered, that a government contemplating Apartheid laws can be convinced that they are unjust, or that a rich local citizen can be found to pay for toxic waste cleanup, are important aspects of a complex structure representing either the international or the domestic systems.

Many of the outcomes in social studies are related to attitudes as well as to cognitions. The approach to integrating attitudes into these definitions of complex cognitive structure has been explored in a preliminary form. It build on studies of social psychologists such as Tourangeau (1987). He proposed that faced with a Likert-attitude item the respondent first reads and understands the question, and then calls to mind the relevant facts or beliefs (often by reference to a cognitive structure or script). For example, for one respondent reference to an issue like welfare may activate a set of interrelated images and feelings involving fraud and Welfare Queens. He noted that for another, the term welfare may evoke a script involving people down on their luck as a result of injustice. Differences in these structures around which knowledge and beliefs are organized result in different answers to the attitude item. In the research described in this paper, this means that attitudes are attached to schemata with different patterns of actors and actions. For some of the adolescents, feelings of outrage appeared to be associated with beliefs about what would happen to the racial minority in the event of Apartheid laws. For others, one could ascertain attitudes toward international debt and its intractability from responses indicating that constraints were likely to overwhelm any possibilitites for success of the proposed actions.

Research on Adolescents: Written Free Responses

In addition to these think-aloud problem solving interviews, participants were also administered a written questionnaire pre- and post-simulation. They were asked to list four questions they would ask the Mexican ambassador to the U.S. and the French ambassador to the U.S. if they had the opportunity to interview them. The purpose was to elicit scripts for diplomacy. The criteria of pre-simulation to post-simulation change also proved useful in analyzing these questions to develop a description of what is meant by a complex cognitive structure.

Complexity Characteristic 3: A student with a complex script as a structure for understanding international diplomacy can see the relevance of potential actions of leaders for dealing with a variety of world issues.

Not only were more questions listed by participants after the simulation than before, but the nature of the questions was different. Pre-simulation several students spoke about the ambassador's personal experience -- did he like the U.S., how long was he likely to stay. Some wanted to know "why are the French so rude to Americans," and another asked about the Bicentennial celebration. In the questions proposed pre-simulation to the Mexican ambassador immigration and drug-smuggling were raised, as well as questions like "why are you so poor," "how is your economy", or "what kind of aid do you want

from the U.S."? The adolescents' scripts for these diplomats seemed to come from recent newspaper headlines. After the simulation, the majority of students seemed to connect a variety of important world issues to the diplomatic scripts for leaders from all countries. They proposed asking both French and Mexican ambassadors something about the Nuclear Proliferation Treaty (why they signed or didn't, how other countries could be encouraged to sign) and about measures such as embargoes which might be taken against Apartheid. These topics were mentioned rarely pre-simulation.

Some specific issues also became especially important for specific countries. After the simulation a large proportion of participants proposed in-depth questions about debt to the Mexican ambassador -- possibility of default and what a moratorium might accomplish or why it would be undesirable. With respect to the French ambassador, substantial numbers also proposed asking him about debt, in particular France's view of a moratorium or what plans they had to contribute to debt reduction in LDC's.

The "propose a question" technique appears to get at students' underlying scripts for what diplomacy is and how political leaders' opinions and actions are relevant to world issues. The technique also seems to assess the acquisition of specific, detailed and connected knowledge about a topic and about related terms (e.g. what moratorium means). The "propose a question" technique has the additional advantage of being suitable for administration in a written form and of being usable across a number of content areas. For example, students could be asked to propose questions for a character in history in order to understand an historical event. These questions could be coded according to the specificity, detail, and connectedness in the associated scripts and the inclusion of relevant conceptual knowledge.

Another type of question was used and related to a fourth complexity characteristic. Students were asked to rank countries and to explain the reasons for those rankings. For example, they were asked to rank countries according to their level of economic development. They were then asked to give reasons for the ranking, and it was those reasons not the specific ranks that were examined. The following definition of complexity resulted.

Complexity Characteristic 4: <u>A student with a complex structure relating to a domestic or international political situation is able to rank or categorize countries or other definable groups along a complex set of dimensions.</u>

Some students gave responses indicating low complexity in the dimensions, reflecting relatively rudimentary types of distinctions and categorizations. For example, in answering a

question about how to differentiate countries with high levels of economic development from those with low levels, some students said that that less developed countries did not have good economies, or that people there were poor or unemployed. They appeared to be using a poor to rich dimension for their ranking, a simple almost tautological answer. The more complex ways of structuring or ranking countries by level of development used by other adolescents included strength of industries, use of technology, high GNP, diversification of exports, absence of debt interfering with economic growth, stability of an economy (without inflation), stability of currency internationally, and ability to influence other countries' economic decisions or organizations like the IMF.

The purpose here is not to indicate that some indices of economic development ought to be taught in preference to others. Rather it is to indicate that the student who sees economic development involving characteristics such as industrialization, freedom from debt, or stable currency has a more complex cognitive structure relating to economic development than the one who says that a country can be judged underdeveloped if poor people live there.

Measures of Complexity in Cognitive Structures as Authentic Assessments in Social Studies

Preliminary data were presented on measures of complexity of underlying cognitive structures derived from oral and written answers to problems, question-generation, and ranking rationales. These represent an attempt to provide an alternative to multiple choice testing built on research on cognitive structures. No new methodology tried on a small sample of adolescents can make valid claims to providing the answer to the need for authentic assessment in social studies. No one would argue that it will be easy or inexpensive to develop and validate on a large-scale basis measures such as these relating to cognitive structures.

This paper, however, represents a starting point by defining what is meant by complexity of cognitive structures with respect to the economic and political system. Four characteristics of complexity in the international political and economic views of young people have been explored, suggesting in each case new indicators or measures. Important next steps are to regularize these into technically more feasible measures and to validate them. These assessment tools have several promising characteristics. They relate to cognitive structures -- schemata of the world or scripts of political leadership -- and this connects them to a rapidly growing body of research and theory on the application of cognitive psychology in the subject matter domains. In other words, they represent higher order knowledge.

These measures have the potential for a high degree of

aut ticity when considered in light of Wiggins' criteria of Int lectual Design Features and Standards of Grading and Scoring. They represent complex intellectual challenges which require the respondent to actively draw upon a structure of knowledge and assess a variety of alternative actions. They stimulate, as he suggested, responses which cannot be divided into right and wrong solutions and encourage students to clarify problems and to elaborate their thinking. Performance on measures such as these should predict accessibility of the organized conceptual knowledge and cognitive structures young people need in their everyday lives as citizens and members of society -- structures that allow them to consider alternative actions when faced with social, economic, or political problems; to judge political leaders, present or past, according to their positions on a wide range of issues; and to view countries in the world on a complex set of dimensions. Students and teachers could be told what high performance consists of on tasks such as these -- it involves generating and evaluating alternative actors and actions.

These techniques also are promising in relation to Wiggins' criterion of Fairness and Equity for authentic assessment and to Miller-Jones' suggestions about tests on which performance reveals cognitive processes. Authentic assessment allows students to demonstrate their own style of solving a problem. This applies to equity by gender as well as equity relating to other characteristics. Several pieces of recent research have demonstrated the lower scores of girls on multiple choice tests in history, social studies, and civic education (Finn & Ravitch, 1987; Anderson et al., 1990; Hammack, et al., 1990; Torney-Purta, 1983; Torney, Oppenheim, and Farnen, 1975; Kneedler,). An analysis of data from eighteen of the students in the Maryland study also indicated that on a short multiple choice test of in international studies, girls did somewhat more poorly than boys. However, from the data analyzed so far it does not appear that girls in the Maryland study have less complex schemata of the world than boys. When encouraged to explore alternatives and consider constraints in a think aloud situation, as many girls as boys described a variety of actors, actions and constraints. In other words, there appear to be few differences between genders in these measures of cognitive complexity of social studies concepts. Although not specifically designed with this in mind, these think aloud techniques may in fact meet needs for gender equity in social studies testing.

In conclusion, the field of performance assessment is expanding rapidly. Innovative techniques are being explored, and strides have been made especially in science. In other fields such as social studies we are at least making a start.

References

Anderson, L. F., Jenkins, L, Leming, J., MacDonald, W., Mullis, I., Turner, M.J, & Wooster, J. (1990) The civics report card: Trends in achievement . Princeton, N. J.: NAEP/ETS.

Armbruster, B. (1984) The problem of "inconsiderate text." In G. Duffy, L. Roehler, & J. Mason (Eds.), Comprehension instruction. New York: Longman. (pp. 202-17).

Ennis, R. (1989) Critical thinking and subject specificity: Clarification and needed research. Educational Researcher, 18, 4-10.

Finn, C. & Ravitch, D. (1987) What do our 17-year olds know? New York: Harper and Row.

Hammack, D.; Hartoonian, M.; Howe, J; Jenkins, L.; Levstik, L.; MacDonald, W.; Mullis, I.; and Owen, E. (1990) The U.S. history report card. Princeton, N. J.: ETS/NAEP.

Hayes, J.R. (1981) The complete problem solver. Philadelphia: Franklin Institute Press.

Kneedler, P. (1988) Differences between boys and girls on California's new statewide assessment in history-social science. Social Studies Review, 27, 96-124.

Miller-Jones, D. (1989) Culture and testing. American Psychologist, 44, 360-66.

Norris, S. (1989) Can we test validly for critical thinking. Educational Researcher, 9, 21-26.

Prawatt, R. (1989) Promoting access to knowledge: Knowledge, strategy, and disposition in students: A research synthesis. Review of Educational Research, 59, 1-41.

Stodolsky, S. (1988) The subject matters: Classroom activity in math and social studies. Chicago: University of Chicago Press.

Torney, J.V.; Oppenheim, A.N.; and Farnen, R. F. (1975) Civic education in ten countries: An empirical study. New York: Wiley.

Torney-Purta, J. (1989) Political cognition and its restructuring in young people. Human Development, 32, 14-23.

Torney-Purta, J. (1990) From attitudes and knowledge to schemata: Expanding the outcomes of political socialization research. In O. Ichilov (Ed.), Political socialization, citizenship education, and democracy. New York: Teachers College (pp. 98-115).

Tourangeau, R. (1987) Attitude measurement: A cognitive perspective. In H. Hippler, N. Schwarz, & S. Sudman (eds.), Social information processing and survey methodology. New York: Springer-Verlag. (pp. 149-162).

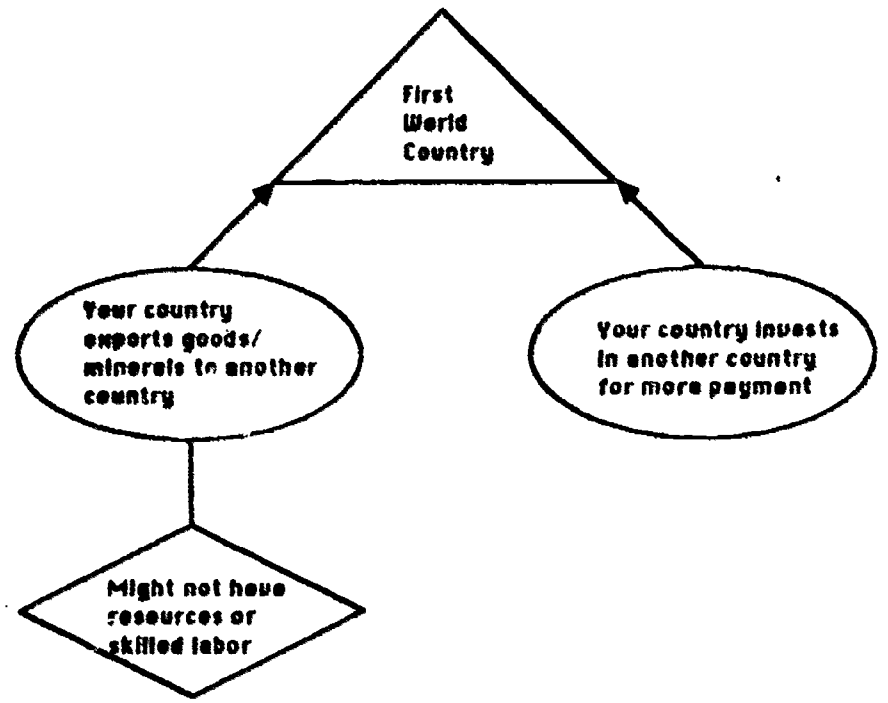
Voss, J.F.; Tyler, S. & Yengo, L. (1983) Individual differences in the solving of social science problems. In R. Dillon & R. Schmeck (Eds.) Individual differences in problem solving (pp. 205-32). New York: Academic Press.

Wiggins, G. (1989) A true test: Toward more authentic and equitable assessment. Phi Delta Kappan, 70, 703-713.

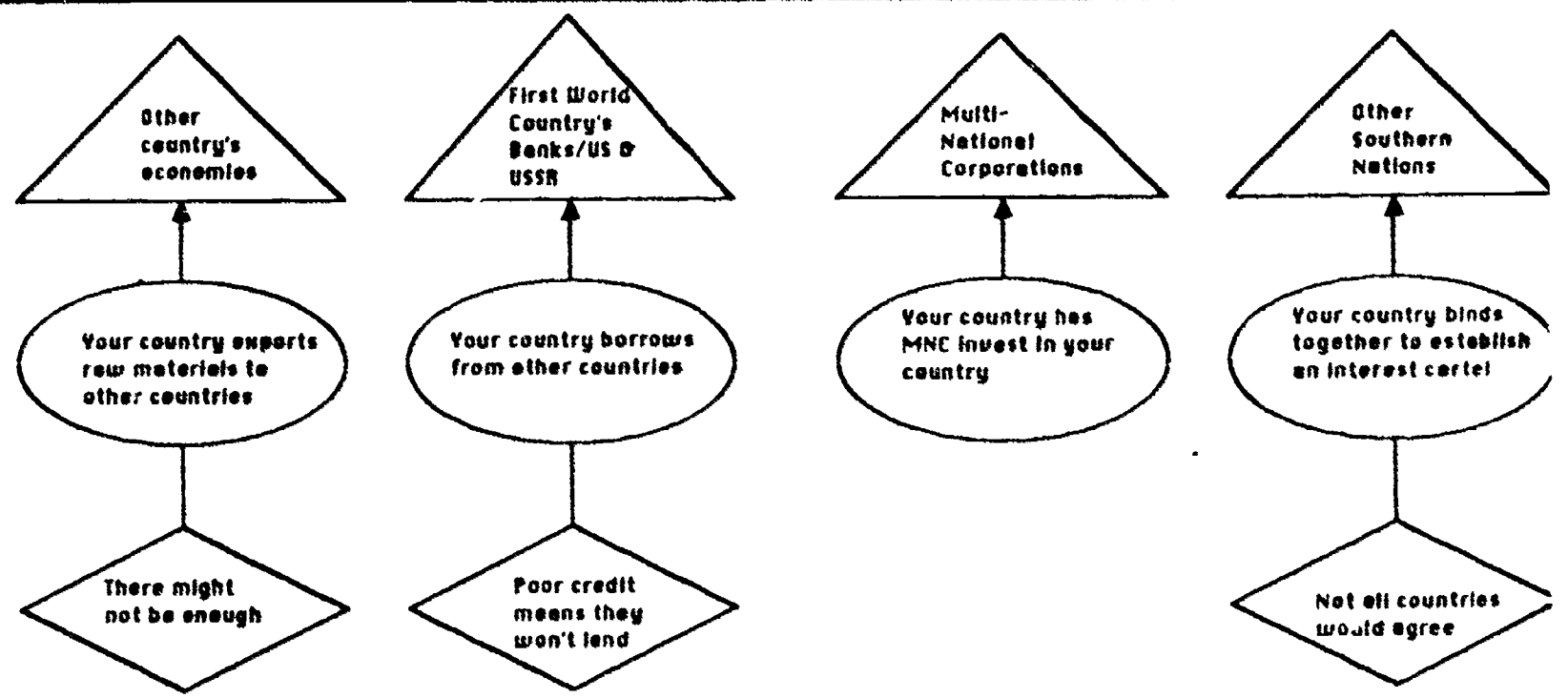
Figure 1

Schemata of Actors, Actions, and Constraints in International Debt Crisis
Pre-Session on Top / Post-Session on Bottom

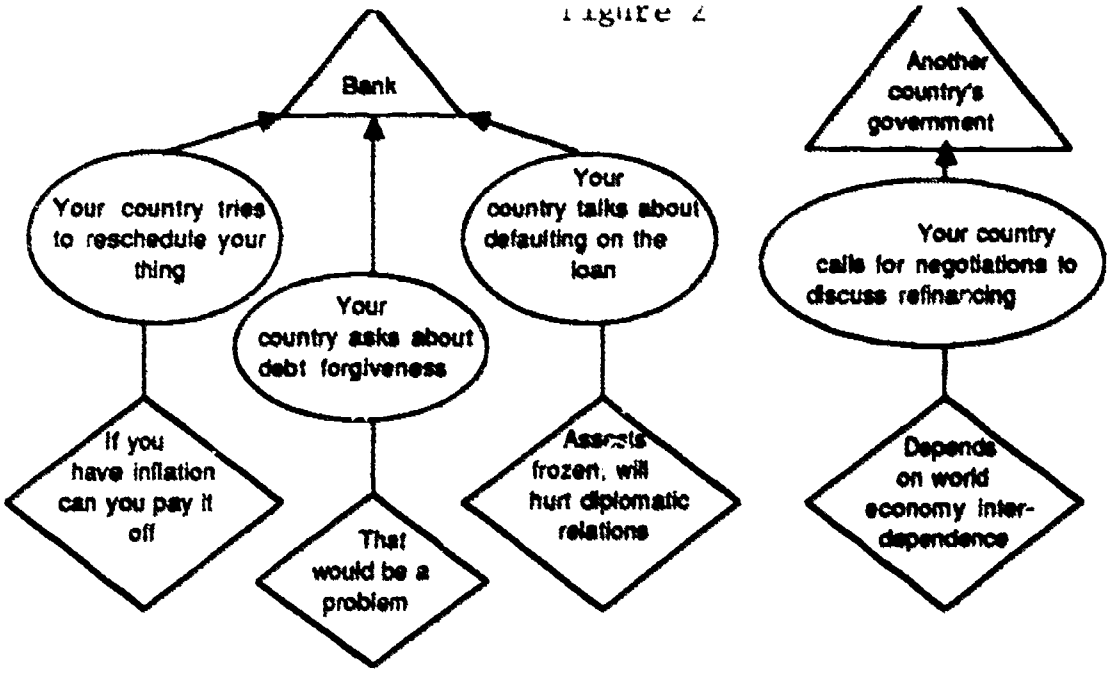
PRE



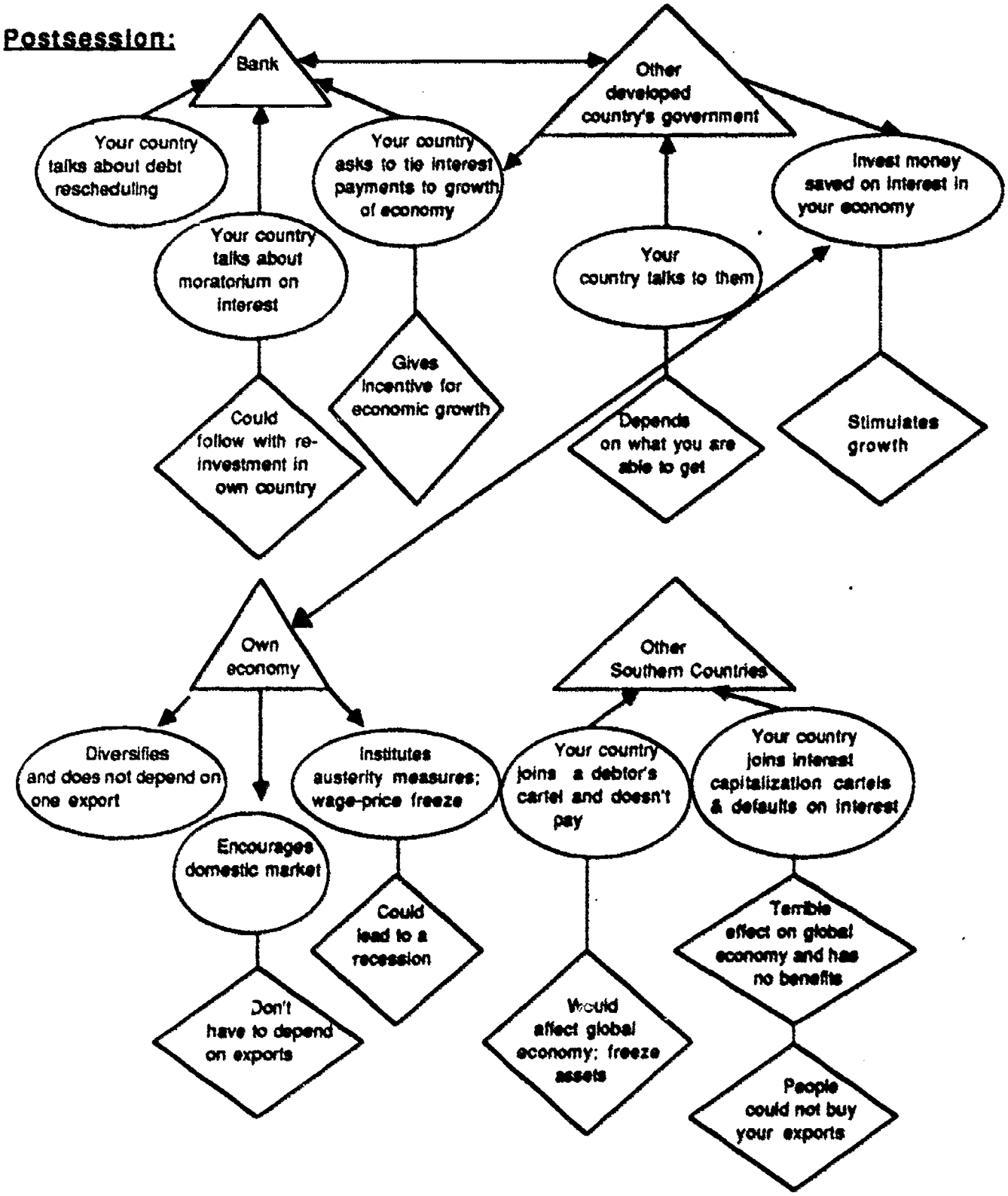
POST



Presession:



Postsession:

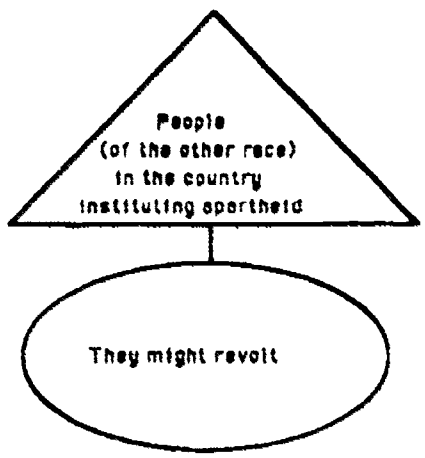


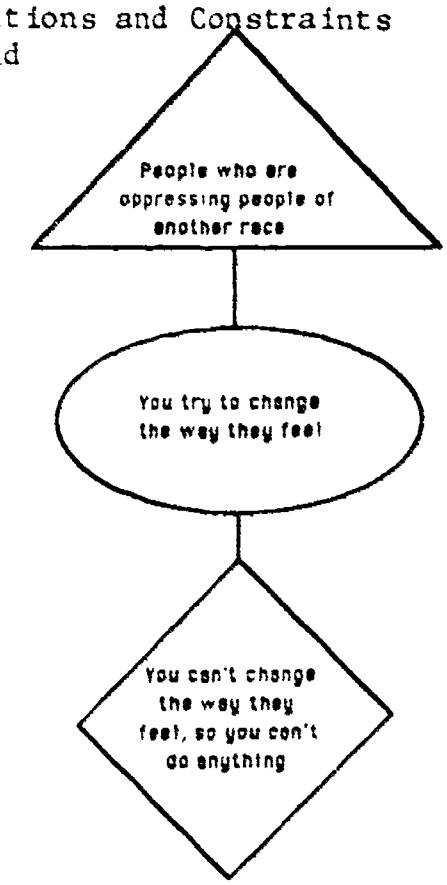
France, female,

Figure 3

Schemata of Actors, Actions and Constraints in Apartheid

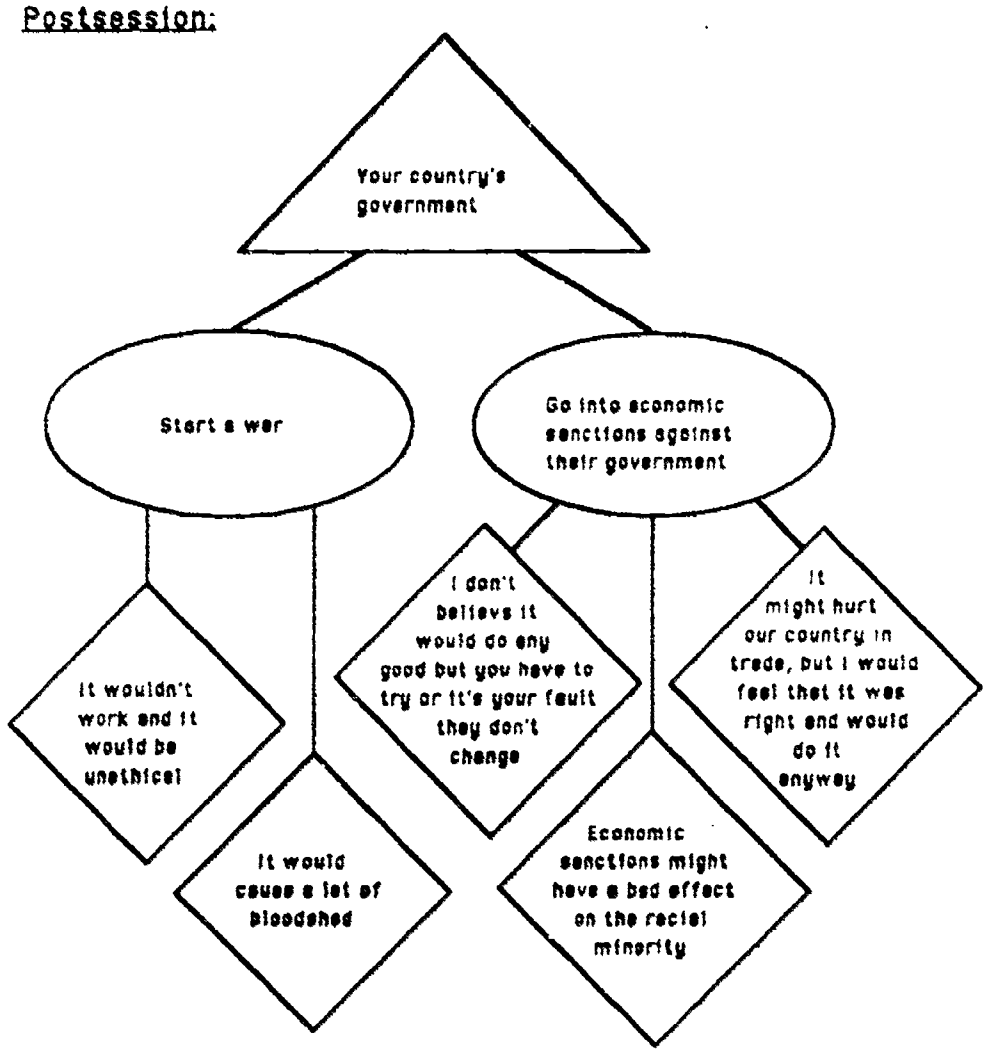
<u>Presession:</u>

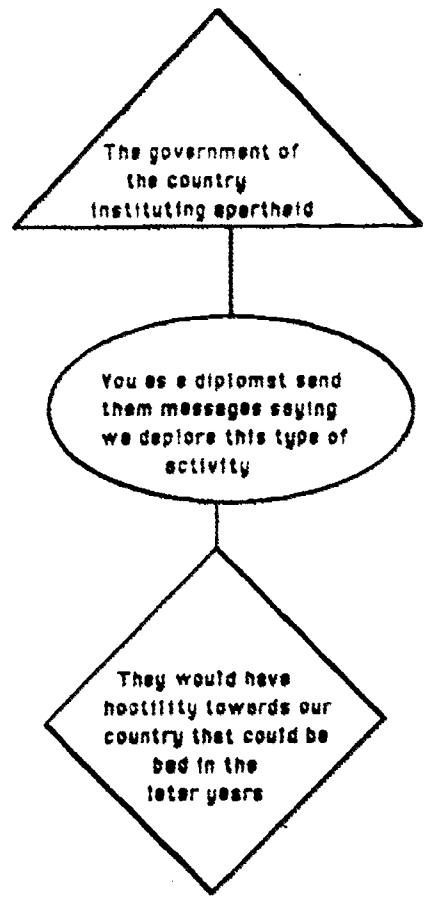




First response was "I don't think there is anything you can do"

<u>Postsession:</u>

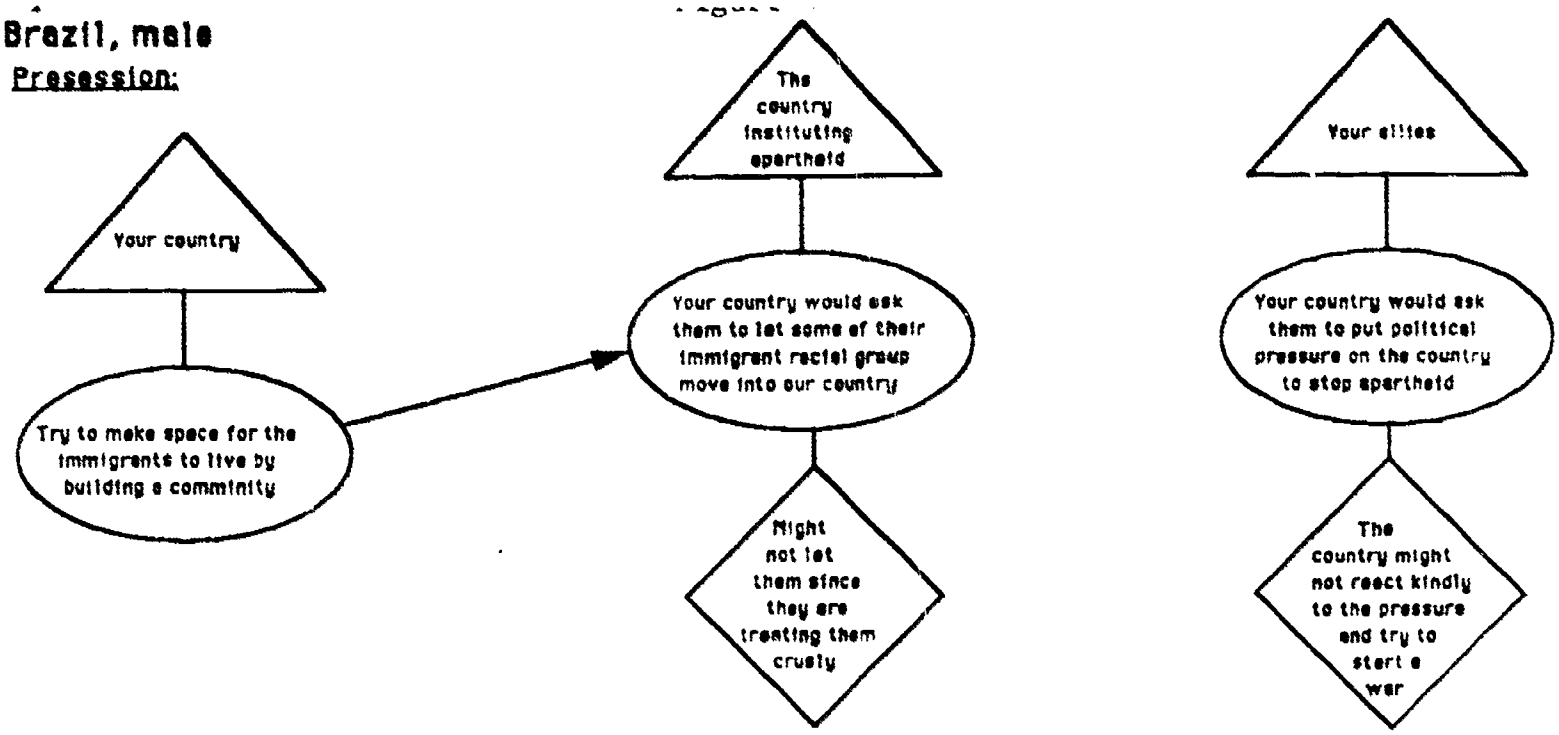




First responses was "Apartheid is very very hard, if not impossible to solve because you can't change what people think and what people feel"

Brazil, male

Presession:



Postsession: